Semántica de actualización [1]

FRANK VELTMAN
(Universidad de Amsterdam)

§ 1. El marco

La definición estándar de validez lógica dice esto: Un argumento es válido si sus premisas no pueden ser todas verdaderas sin que su conclusión sea también verdadera. La amplia mayoría de las teorías desarrolladas hasta el momento han adoptado esta definición de validez como su punto de partida. En consecuencia, el núcleo de tales teorías consiste en una especificación de condiciones de verdad.

El núcleo de la teoría presentada más abajo no consiste en una especificación de condiciones de verdad, sino de condiciones de actualización. De acuerdo con esta teoría, el eslogan "Se conoce el significado de una oración si se conocen las condiciones bajo las cuales es verdadera" debería sustituirse por este otro: Se conoce el significado de una oración si se conoce el cambio que induce en el estado epistémico de quien acepte las noticias que transmite [2]. De esta forma, significado se convierte en una noción dinámica: el significado $[\phi]$ de una oración ϕ es una operación sobre estados epistémicos o estados de información, como yo prefiero llamarlos [3].

[1] Este artículo contiene la traducción española de las dos primeras secciones de Veltman [1991], un artículo mucho más largo que será publicado en el *Journal of Philosophical Logic*. Quiero agradecer al profesor Acero su oferta de hacer esta traducción y el haber señalado algunas oscuridades del original.

[2] Esta concepción del significado subyace a mucha labor reciente en semántica formal. Su origen puede retrotraerse al trabajo de Robert Stalnaker sobre presuposición y aserción. (Véase, por ejemplo, Stalnaker [1979].) Adoptó otra forma en la obra de Hans Kamp e Irene Heim sobre anáfora y en la de Peter Gärdenfors sobre la dinámica de la creencia. (Véase, por ejemplo, Kamp [1981], Heim [1982] y Gärdenfors [1984].)

[3] Diferentes aplicaciones del marco darán lugar a diferentes aplicaciones de la noción de estado de información. En la segunda sección de este artículo la única cosa que se tiene en cuenta es el conocimiento que pueda tener un agente y, consiguientemente, identificamos su estado de información con lo que él o ella sabe. Sin embargo, en otras aplicaciones otras cosas podrían también importar: las expectativas del agente, sus deseos, etc.

Revista de Filosofía. 3.ª época, vol. IV (1991), núm. 6, págs. 271-284. Editorial Complutense. Madrid

Sea σ un estado de información y ϕ una oración con significado $[\phi]$. Escribo $\sigma[\phi]$ para representar el estado de información que resulta de actualizar σ con ϕ [4]. En la mayoría de los casos $\sigma[\phi]$ será diferente de σ. Entonces la información transmitida por ϕ se halla ya subsumida en σ: si σ se actualiza con ϕ, el estado de información resultante resulta ser σ de nuevo. En tal caso, es decir, cuando $\sigma[\phi] = \sigma$, escribimos $\sigma \Vdash \phi$ y decimos que ϕ *es aceptada en* σ.

Dos explicaciones de validez lógica se sugieren entonces por sí solas: (i) Un argumento es válido$_1$ si uno no puede aceptar todas sus premisas sin tener que aceptar también su conclusión. Más formalmente:

$$\psi_1,...,\psi_n \models_1 \phi \text{ syss } \sigma \Vdash \phi \text{ para toda } \sigma \text{ tal que } \sigma \Vdash \psi_i$$

(para todo $1 \leq i \leq n$).

(ii) Un argumento es válido$_2$ syss la actualización de un estado de información σ con las premisas $\psi_1,...,\psi_n$, en ese orden, da lugar a un estado de información en el que la conclusión ϕ es aceptada. Más formalmente:

$$\psi_1,...,\psi_n \models_2 \phi \text{ syss para todo } \sigma, \sigma[\psi_1] ... \sigma[\psi_n] \Vdash \phi.$$

Es fácil ver que los argumentos que son válidos$_2$ también son válidos$_1$. La conversa no es verdadera.

1.1. PROPOSICIÓN

Considérese el siguiente principio:

Indempotencia: Para cualquier estado de información σ y oración ϕ,

$$\sigma[\phi] \Vdash \phi$$

Dado este principio, los siguientes son equivalentes:

(a) Si $\psi_1,...,\psi_n \models_1 \phi$, entonces $\psi_1,...,\psi_n \models_2 \phi$;
(b) *Estabilidad:* Si $\sigma \Vdash \phi$, entonces $\sigma[\psi_1] ... [\psi_n] \Vdash \phi$;
(c) *Monotonicidad a la derecha:* Si $\psi_1,..., \psi_n \Vdash_2 \phi$, entonces $\psi_1,..., \psi_n, \chi_1,..., \chi_k \Vdash_2 \phi$.

En lo que sigue, daremos por cierto el principio de *Idempotencia* [5] ¿qué significaría "actualizar su información con ϕ" si no, al menos, "cambiar su

[4] Puesto que σ es el argumento y $[\phi]$ la función, habría estado más en línea con la práctica común escribir '$[\phi](\sigma)$'. La notación presente resulta más conveniente a la hora de tratar con textos. Ahora podemos escribir '$\sigma[\psi_1] ... [\psi_n]$ —en lugar de "$[\psi_n](... [\psi_1](\sigma) ...)$'— para representar el resultado de actualizar σ con la sucesión de oraciones $\psi_1,..., \psi_n$.

[5] En realidad, hay muchas oraciones del lenguaje natural respecto de las cuales no

información de tal modo que Usted llegue a aceptar ϕ"? Sin embargo, no asumiremos el principio de *Estabilidad*. Hay oraciones que no son estables. A menudo se las aceptará en un primer momento por disponer de información limitada, pero sólo para rechazarlas cuando se tenga más información a mano.

EJEMPLO

Los ejemplos más claros de oraciones inestables han de encontrarse entre las oraciones en las que aparezcan cualificaciones modales como "supuestamente", "probablemente", "debe", "puede" o "podría". Que semejantes oraciones no son estables lo muestran las dos siguientes sucesiones de oraciones. El procesamiento de la primera no origina ningún problema, mientras que el de la segunda sí que lo hace:

— Alguien llama a la puerta... Debe de ser Juan... Es María.

— Alguien llama a la puerta... Es María... Debe de ser Juan.

Estas dos sucesiones contienen las mismas oraciones. Sólo el orden cambia. No obstante, la primera sucesión tiene sentido, pero no la segunda. Explicación: Es normal que nuestras propias expectativas sean refutadas por los hechos, eso es lo que refleja la primera sucesión. Pero si uno sabe ya algo, es un poco tonto pretender que todavía se espera algo más, que es lo que sucede en la segunda.

Una de las ventajas del enfoque dinámico es que estas diferencias pueden explicarse. El marco nos capacita para habérnoslas con sucesiones de oraciones, con textos completos. Sea ϕ_1 = "Alguien llama a la puerta", ϕ_2 = "Debe de ser Juan" y ϕ_3 = "Es María". Para cualquier estado de información σ, cabría comparar $\sigma\,[\phi_1]\,[\phi_2]\,[\phi_3]$ con $\sigma\,[\phi_1]\,[\phi_2]\,[\phi_3]$ y ver si hay diferencias entre ambas actualizaciones.

La Proposición 1.1 dice que en ausencia de *Estabilidad* las dos explicaciones de validez darán lugar a lógicas diferentes. Nótese que la primera noción de validez es monotónica. Si un argumento con premisas suscritas 1,...,n y conclusión ϕ es válido $_1$, entonces seguirá siendo válido $_1$ si se añaden más premisas a $\psi_1,..., \psi_n$. Dada la proposición 1.1, la segunda noción será nomotónica. Nótese, sin embargo, que la segunda noción es al menos monotónica a la izquierda:

$$\text{Si } \psi_1,..., \psi_n \models_2 \phi, \text{ entonces } \chi_1,..., \chi_k, \psi_1,..., \psi_n \models_2 \phi.$$

Lo que falla es la monotonicidad a la derecha:

$$\text{Si } \psi_1,..., \psi_n \models_2 \phi, \text{ entonces } \psi_{,}..., \psi_n, \chi_1,..., \chi_k \not\models_2 \phi.$$

vale el Principio de Idempotencia. Las oraciones con pronombres anafóricos y oraciones paradójicas como el Mentiroso pueden servir de ejemplos. Los lenguajes formales discutidos en este escrito no contienen oraciones como ésas. Véase Groenendijk & Stokhof [1989] para más información sobre anáfora, y Groeneveld [1989] para un análisis dinámico de la Paradoja del Mentiroso.

Hay una tercera explicación de validez, una que es de vital importancia en el razonamiento por defecto: Un argumento es válido $_3$ syss la actualización del mínimo estado de información "**1**" con las premisas $\psi_1,..., \psi_n$, en ese orden, induce un estado de información σ en el que se acepta la conclusión ϕ. Formalmente:

$$\psi_1,..., \psi_n \models_3 \phi \text{ syss } \mathbf{1}\,[\psi_1] \ldots [\psi_n] \Vdash \phi.$$

Naturalmente, esta definición presupone que existe una cosa como el mínimo estado de información.

Al igual que la validez $_1$, la validez $_3$ da lugar a una lógica más fuerte que la de la validez $_2$. Si $\psi_1,..., \psi_n \models_2 \phi$, entonces $\psi_1,..., \psi_n \models_3 \phi$, pero la conversa no vale. Al igual que la validez $_2$, la validez $_3$ no es monotónica a la derecha. Pero validez $_3$ tampoco es monotónica a la izquierda.

EJEMPLO

En Veltman [1991] se presenta una lógica de principios vigentes por defecto [6] de acuerdo con la cual el siguiente esquema argumentativo es válido $_3$:

Premisa 1:	Los P normalmente son R
Premisa 2:	x es P
Conclusión:	Supuestamente, x es R

Este argumento permanece válido $_3$ si uno aprende más cosas sobre el objeto x en la medida en que no haya evidencia alguna de que la nueva información sea pertinente para con la conclusión. También en el caso siguiente la interferencia vale:

Premisa 1:	Los P normalmente son R
Premisa 2:	x es P y x es Q
Conclusión:	Supuestamente, x es R

Sin embargo, si se añade la regla "Los Q normalmente no son R" encima de las premisas 1 y 2, el argumento resultante ya no es válido $_3$. Es decir, si todo lo que sabemos es

Premisa 1:	Los Q normalmente no son R
Premisa 2:	Los P normalmente son R
Premisa 3:	x es P y x es Q

[6] Así vierto la expresión "*logic for default reasoning*". Consiguientemente, los *default principles* serán principios vigentes por defecto. "*Default reasoning*" se traduce simplemente por "razonamiento por defecto". (Nota del traductor.)

entonces permanece abierta la cuestión de si podemos suponer que *x* sea *R*. Claramente, el objeto *x* debe ser una excepción a alguna de las reglas, pero no hay ninguna razón para esperar que sea una excepción a una regla mejor que a la otra.

La adición de otros principios vigentes por defecto puede equilibrar la balanza. Por ejemplo, si añadiéramos "Los *Q* son normalmente *P*" como premisa, obtendríamos el siguiente argumento válido $_3$:

Premisa 1:	Los *Q* normalmente son *P*
Premisa 2:	Lo *Q* normalmente no son *R*
Premisa 3:	Los *P* normalmente son *R*
Premisa 4:	*x* es *P* y *x* es *Q*
Conclusión:	Supuestamente, *x* no es *R*

En presencia del principio "Los *Q* normalmente son *P*", el principio "Los *Q* normalmente no son *R*" adquiere prioridad sobre el principio "Los *P* normalmente son *R*". (Si se desea un ejemplo concreto, tómese '*x* es *P*' como "es adulto", "*x* es *Q*" como "*x* es estudiante" y "*x* es *R*" como "*x* es alguien que tiene un puesto de trabajo").

Unos cuantos comentarios. Un respecto en el que esta teoría del razonamiento por defecto difiere de otras teorías en curso es que la distinción entre conclusiones cancelables y no-cancelables se hace manifiesta al nivel del lenguaje-objeto: No es válido$_3$ concluir de "Los *P* normalmente son *R*" y "*x* es *R*" que *x sea R*; sólo lo es concluir que *supuestamente x* es *R*. Esta cualificación explicita el hecho de que se ha extraído una conclusión inestable y, por lo tanto, cancelable. En otras teorías uno puede inferir bien "*x* es *R*" de las premisas, pero se considera que se está ante un caso especial de inferencia. Allí donde estas otras teorías consideran el razonamiento por defecto una clase especial de razonamiento con oraciones comunes, nosotros preferimos concebirlo como un tipo común de razonamiento con una clase especial de oraciones.

Una segunda diferencia con respecto a muchas teorías es que las cuestiones de prioridad, que probablemente surgen en el caso de supuestos conflictivos (cf. los dos últimos ejemplos de más arriba), se deciden en el nivel de la semántica. Que "Los *Q* normalmente no son *R*" se impone a "Los *P* normalmente son *R*" en presencia de "Los *Q* normalmente son *P*" es algo que refuerza lo que estos principios *significan*. No es algo que haya de estipularse además de la semántica —como muchas teorías dirían—, sino algo que ésta explica.

En tercer lugar: Ninguno de los argumentos discutidos más arriba son válidos $_1$ o válidos $_2$. Tanto la definición de validez $_1$ como la de validez $_2$ contienen una cuantificación sobre todos los posibles estados de información. Así, pues, debemos contar con estados de información en los que se *sabe* que el objeto bajo consideración es anormal en los aspectos pertinentes.

El siguiente argumento, sin embargo, es válido $_i$, para $i = 1, 2, 3$:

Premisa 1:	Los P normalmente son R
Premisa 2:	x es P
Premisa 3:	x puede ser R
Conclusión:	Supuestamente, x es R

Aquí, el "puede" indica que, dado todo lo que uno sabe, no se dan circunstancias anormales.

Observación final: Es fácil verificar que la validez $_3$ y la validez $_2$ se ajustan al principio de *Monotonicidad Restringida:*

$$\text{Si } \psi_1,\ldots, \psi_n \models_i \phi \text{ y } \psi_1,\ldots, \psi_n \models_i \chi,$$

entonces

$$\psi_1,\ldots, \psi_n, \chi \models_i \phi \ (i = 2, 3)$$

Más aún, ambos cumplen con el principio de *Generación de Lemas:*

$$\text{Si } \psi_1,\ldots, \psi_n \models_i \chi \text{ y } \psi_1,\ldots, \psi_n, \chi \models_i \phi,$$

entonces

$$\psi_1,\ldots, \psi_n \models_i \phi \ (i = 2,3)$$

En la actualidad, Monotonicidad Restringida, Corte y Reflexividad,

$$\phi \models \phi \quad (i = 2,3)$$

se consideran frecuentemente las condiciones mínimas que debe satisfacer cualquier relación de consecuencia razonable. (Véase Makinson [1989].)

§ 2. Un ejemplo: *podría*

2.1. DEFINICION

Sea $\mathcal{A}$ un conjunto finito de *oraciones atómicas*. Asociemos a $\mathcal{A}$ dos lenguajes, $L_0^{\mathcal{A}}$ y $L_1^{\mathcal{A}}$. Ambos tienen a $\mathcal{A}$ como vocabulario no-lógico. L_0 tiene como vocabulario lógico exclusivo un operador monario $\neg$, dos operadores binarios $\wedge$ y $\vee$ y dos paréntesis, (y). La oraciones de $L_0^{\mathcal{A}}$ son precisamente las que uno esperaría para un lenguaje con semejante vocabulario.

$L_1^{\mathcal{A}}$ tiene en su vocabulario lógico un operador monario adicional: *podría*, y un signo de puntuación denotado mediante ";". Una cadena ϕ de símbolos es una oración de $L_1^{\mathcal{A}}$ syss hay alguna oración ψ de L_0 tal que o bien $\phi = \psi$ o bien $\phi =$ *podría* ψ.

El conjunto de los textos $L_1^{\mathcal{A}}$ es el más pequeño que satisface las condiciones siguientes:

(i) Si ϕ es una oración de $L_1^{\mathcal{A}}$, entonces ϕ es un texto de $L_1^{\mathcal{A}}$;

(ii) Si ν es un texto de $L_1^{\mathcal{A}}$ y ϕ es una oración de $L_1^{\mathcal{A}}$, entonces $\nu ; \phi$ es un texto de $L_1^{\mathcal{A}}$.

Más abajo, "p", "q", "r", "p_1", "q_1", "r_1", "p_2", "q_2", etc. se usan como metavariables de oraciones atómicas. Diferentes metavariables refieren a *diferentes* oraciones atómicas. Las letras griegas ϕ, ψ, y χ se usan como metavariables de oraciones arbitrarias.

La idea que se halla tras el análisis de *podría* que se ofrece más abajo es ésta: Uno tiene que estar de acuerdo con la información de que *podría* ϕ si ϕ es consistente con lo que él sabe, o mejor, con lo que uno piensa que es lo que sabe. En caso contrario, *podría* ϕ ha de rechazarse.

A fin de encajar esta idea en un modelo matemático, necesitamos un modo de representar el saber del agente. Más abajo se representa un estado de conocimiento [7] σ de un agente por medio de un conjunto de subconjuntos de $\mathcal{A}$. Intuitivamente, un subconjunto i de $\mathcal{A}$ será un elemento de σ si, para todo lo que el agente sabe en el estado σ, este conjunto *podría* dar una imagen correcta de la realidad. De forma más precisa, para todo lo que sabe el agente en un estado σ, no se excluye la posibilidad de que las oraciones atómicas en i son todas verdaderas y el resto falsas. El conjunto potencia de $\mathcal{A}$ determina el espacio de posibilidades *a priori:* si acontece que el agente no sepa nada en absoluto, entonces cualquier subconjunto de $\mathcal{A}$ podría representar la realidad correctamente. Cuando el saber del agente aumente, σ se encogerá hasta que conste de un único subconjunto de $\mathcal{A}$. Entonces el saber del agente es completo. Así, pues, el crecimiento del saber se entiende como un proceso de eliminación.

Para un agente que no sepa nada en absoluto, cualquier subconjunto de $\mathcal{A}$ podría representar la realidad correctamente –semejante conjunto representaría una "forma en que podría ser el mundo". A falta de un término mejor, de aquí en adelante me referiré a los subconjuntos de $\mathcal{A}$ llamándolos mundos posibles [8].

---

[7] Uso la frase "conocimiento" y "estado de conocimiento" donde algunos lectores podrían preferir "creencias" y "estados de creencia". De hecho, quiero que los estados de información σ representen algo entre medias: si σ es el estado de un agente dado, habría de denotar lo que el agente considera que es lo que él sabe. Las cosas de las que el agente diría que simplemente cree no contarían. Pero muy bien podría ocurrir que algo que el agente juzga "conocido" sea de hecho falso.

[8] Esto es un poco equívoco, puesto que la frase "mundo posible" sugiere que no estamos hablando acerca de un modo en que las cosas podrían ser, sino de algo que es de ese modo. Cf. Stalnaker [1976]. En segundo lugar, resultará claro que los "mundos posibles" implicados no tiene por qué poseer la complejidad de una completa alternativa

2.2. DEFINICION

Sea $\mathscr{A}$ el conjunto de las oraciones atómicas y W el conjunto potencia de $\mathscr{A}$. Entonces

(i) σ es un *estado de información* syss $\sigma \subseteq W$;
(ii) Sean σ y τ estados de información. Entonces σ *es al menos tan fuerte como* τ syss $\sigma \subseteq \tau$;
(iii) **1**, el estado mínimo, es el estado de información dado por W; **0**, el estado absurdo, es el estado de información dado por el conjunto vacío.

La noción de estado de información es dependiente del lenguaje: diferentes conjuntos de oraciones atómicas dan lugar a diferentes conjuntos de estados de información posibles. La definición oscurece esto. Sería más exacto hablar de estados de $\mathscr{A}$-información y del estado $\mathscr{A}$-mínimo. Más abajo usaré ocasionalmente esta segunda terminología, en particular cuando estemos listos para demostrar que en asuntos de lógica no es importante conocer exactamente qué lenguaje está en juego.

2.3. DEFINICION

Sea $\mathscr{A}$ un conjunto del tipo especificado. Para cada oración ϕ de $L_1^{\mathscr{A}}$ y todo estado de información σ, $\sigma\,[\phi]$ viene dado por la siguiente definición recursiva:

$$
\begin{array}{rl}
\text{Atomos:} & \sigma\,[p] = \sigma \cap \{w \in W \mid p \in w\} \\
\neg\!: & \sigma\,[\neg\,\phi] = \sigma - \sigma\,[\phi] \\
: & \sigma\,[\phi \wedge \psi] = \sigma\,[\phi] \cap \sigma\,[\psi] \\
: & \sigma\,[\phi \vee \psi] = \sigma\,[\phi] \cup \sigma\,[\psi] \\
\textit{podría}\text{:} & \sigma\,[\textit{podría}\ \phi] = \sigma,\ \text{si}\ \sigma\,[\phi] \neq \mathbf{0} \\
 & \sigma\,[\textit{podría}\ \phi] = \mathbf{0},\ \text{si}\ \sigma\,[\phi] \neq \mathbf{0} \\
\text{Textos:} & \sigma\,[\nu;\phi] = \sigma\,[\nu]\,[\phi].
\end{array}
$$

Las cláusulas de actualización dicen, para cada oración ϕ y cada estado de información σ, cómo σ cambia cuando alguien que se halla en el estado σ acepta ϕ. Si $\sigma\,[\phi] \neq \mathbf{0}$, decimos que ϕ *es aceptable* en σ. Si $\sigma\,[\phi] = \mathbf{0}$, que ϕ *no es aceptable* en σ. Estas nociones quieren ser normativas mejor que descriptivas: Si $\sigma\,[\phi] = \mathbf{0}$, un agente en el estado σ no *debería* aceptar ϕ. Y si $\sigma\,[\phi] = \sigma$, un agente que se halle en el estado σ *tiene* que aceptar ϕ. El agente que rehuse obrar de esta guisa está voluntaria o involuntariamente rompiendo las convenciones que gobiernan el uso de $\neg$, $\wedge$, $\vee$, *podría*, etc.

Es también importante tener presente que estas nociones tienen poco o nada que ver con las nociones de verdad y falsedad. Es perfectamente posible

---

al mundo real. Representan un modo en que *algunas* cosas podrían ser, mejor que un modo en que *todas* las cosas podrían ser.

que σ $[p] = \mathbf{0}$, mientras que de hecho p sea verdadera; o que σ $[p] = \sigma$ aunque de hecho p sea falsa.

Supongamos que p es de hecho verdadera y que σ $[p] = \mathbf{0}$. Dada la terminología introducida más arriba, p no es aceptable para un agente en el estado σ. ¿Significa esto que un agente en el estado σ debe rehusarse a aceptar p, incluso cuando él o ella llegue a conocer los hechos? Naturalmente que no. La oración p no es aceptable *en* el estado σ. Así, pues, el agente debe *revisar* σ de tal modo que σ *se haga* aceptable.

La definición anterior no se ocupa para nada de la revisión de estados de información: Las cláusulas de actualización no dicen, para cualquier oración ϕ, cómo debe ser revisado un estado σ en el que ϕ no sea aceptable para que ϕ sea aceptada finalmente. Se detienen en el punto en que resulta claro que surgiría una inconsistencia si se incorporase a σ la información contenida en ϕ.

Nótese que, para cada oración ϕ, $\mathbf{0}\,[\phi] = \mathbf{0}$. Por lo tanto, toda oración se acepta en el estado absurdo, pero ninguna es aceptable en él. Esto explica cómo puede ser que, aun no ocupándonos de la revisión de los estados de información, el principio de *Idempotencia* todavía siga siendo válido. Incluso si una oración ϕ no es aceptable en σ, el resultado de actualizar σ o con ϕ es un estado de información en el que se acepta ϕ.

Aunque no estemos tratando de la revisión de creencias, puede muy bien suceder que en un estadio se acepte una oración y que esa misma oración sea rechazada más tarde: El Principio de Estabilidad falla. La cuestión estriba en que la revisión no es la única fuente de inestabilidad; la comprobación es otra. Las oraciones de la forma *podría* ϕ proporcionan un ejemplo de este otro caso. Según dice la definición, todo lo que cabe hacer cuando se nos dice que podría ser el caso que ϕ es estar o no de acuerdo con ello. Si ϕ es aceptable en un estado de información σ, entonces se ha de aceptar *podría* ϕ. Y si ϕ no es aceptable en σ, tampoco lo es *podría* ϕ.

Es claro, entonces, que las oraciones de la forma *podría* ϕ constituyen una invitación a someter a σ a comprobación mejor que a incorporarle nueva información. Y el resultado de esta prueba puede ser positivo en un estadio y negativo más tarde. En el estado de información mínimo se tiene que aceptar que podría estar lloviendo, pero tan pronto como se sepa que no está lloviendo ha de rechazarse *podría estar lloviendo*.

2.4. DEFINICION

Un texto υ es *consistente* syss hay un estado de información σ tal que σ $[\nu] \neq \mathbf{0}$.

Una vez más, puesto que el conjunto de los estados de información varían con el vocabulario no-lógico del lenguaje en el que el texto ha sido formulado, habría sido más exacto introducir una noción de $\mathcal{A}$-consistencia. Los siguientes lema y proposición muestran, sin embargo, que este prefijo puede omitirse.

2.5. LEMA

Sea $\mathscr{A} \subseteq \mathscr{A}'$. A cada $\mathscr{A}$-estado σ asociamos un $\mathscr{A}'$-estado σ^* dado por medio de

$$\sigma^* = \{j \subseteq \mathscr{A}' \mid j \cap \mathscr{A} \in \sigma\}$$

A cada $\mathscr{A}'$-estado asociamos un $\mathscr{A}$-estado σ^{o} dado por medio de

$$\sigma^{o} = \{j \subseteq \mathscr{A} \mid j = i \cap \mathscr{A}, \text{ para algún } i \in \sigma\}$$

Ahora, para cada ϕ de $L_1^{\mathscr{A}}$ vale lo siguiente:

ia) si σ es un $\mathscr{A}$-estado, entonces $\sigma\,[\phi]^* = \sigma^*\,[\phi]$;
ib) si σ, τ son $\mathscr{A}$-estados y $\sigma \neq \tau$, entonces $\sigma^* \neq \tau^*$;
iia) sin σ es un $\mathscr{A}'$-estado, entonces $\sigma\,[\phi]^{o} = \sigma^{o}\,[\phi]$;
iib) si σ es un $\mathscr{A}'$ estado y $\sigma\,[\phi] \neq \sigma$, entonces $\sigma^{o}\,[\phi] \neq \sigma^{o}$.

En general, no sucede que si σ y τ son $\mathscr{A}'$-estados y $\sigma \neq \tau$, también $\sigma^{o} \neq \tau^{o}$. Afortunadamente, la propiedad especificada en *iib*) es suficientemente fuerte para demostrar la siguiente proposición.

2.6. PROPOSICION

Sean $p_1,..., p_k$ las oraciones atómicas que aparecen en $\psi_1,..., \psi_n, \phi$. Supóngase que $\{p_1,..., p_k\} \subseteq \mathscr{A}$ y $\{p_1,..., p_k\} \subseteq \mathscr{A}'$. Entonces

(*i*) El argumento $\psi_1,..., \psi_n/\therefore\ \phi$ es $\mathscr{A}$-válido$_i$ syss es $\mathscr{A}'$-válido$_i$ ($i = 1, 2, 3$);
(*ii*) $\psi_1;...; \psi_n$ es $\mathscr{A}$-consistente syss $\psi_1;...; \psi_n$ es $\mathscr{A}'$-consistente.

Supongamos que $p_1,..., p_k$ son las oraciones atómicas que aparecen en el argumento $\psi_1;...; \psi_n/\therefore\ \phi$. Dada la proposición 2.6, podemos estar seguros de que la respuesta de si $\psi_1,..., \psi_n/\therefore\ \phi$ es válido$_i$ ($i = 1, 2, 3$) es independiente del lenguaje, tal y como tendría que ser. En realidad, al buscar la respuesta podemos siempre limitarnos a tener en cuenta el conjunto de los estados generados por $\mathscr{A} = \{p_1,..., p_k\}$. Ya que únicamente hay un número finito de ellos, las lógicas generadas por las tres nociones de validez son decidibles.

Los siguientes ejemplos ilustran las afirmaciones hechas en la sección precedente.

EJEMPLOS

(*i*) *podría* $\neg p$; p es conssistente;
p; *podría* $\neg p$ no es consistente;
(Cf. el primer ejemplo de #.1)
(*ii*) *podría* $\neg p \models_2$ *podría* $\neg p$, pero *podría* $\neg p$, $p \models_2$ *podría* $\neg p$;
podría $\neg p \models_3$ *podría* $\neg p$, pero *podría* $\neg p$, $p \models_3$ *podría* $\neg p$;
(Con otras palabras, ni la validez $_2$ ni la validez $_3$ son monotónicas a la derecha.)

(*iii*) $\models_3$ *podría p*, pero $\neg p \not\models_3$ *podría p*.
(Con otras palabras, la validez $_3$ no es monotónica a la izquierda.)

Nótese que la lógica generada por la tercera noción de validez no está cerrada bajo sustitución: $\models_3$ *podría p*, pero $\not\models_3$ *podría* $(p \wedge \neg p)$.

Un estudio sistemático del comportamiento de *podría* con respecto a las tres nociones de validez habrá de dejarse para otra ocasión. Lo que sigue son algunas observaciones preliminares.

2.7. PROPOSICION

Sean σ y τ estados de información y ϕ una oración.

(*i*) $\sigma[\phi][\phi] = \sigma[\phi]$;
(*ii*) $\sigma \subseteq \sigma$;
(*iii*) Si $\sigma \subseteq \tau$, entonces $\sigma[\phi] \subseteq [\phi]$;
(*iv*) Si ϕ es una oración de $L_0^{\mathcal{A}}$, entonces $(\sigma \cup \tau) \subseteq \sigma[\phi] \cup \tau[\phi]$.

EJEMPLO

La cláusula (*iv*) de arriba no vale para $L_1^{\mathcal{A}}$:
Supongamos que $p \in w_1$, $p \in w_2$. Tómese $\sigma = \{w_1\}$ y $\tau = \{w_2\}$.
Entonces $(\sigma \cup \tau)$ [*podría p*] $= \sigma \cup \tau$. Pero σ [*podría p*] $\cup\ \tau$ [*podría p*] $= \sigma \cup \mathbf{0} = \sigma$.

2.8. PROPOSICION

(*i*) (*ii*) y (*iv*) de 2.7 implican *Estabilidad*. Toda oración de $L_0^{\mathcal{A}}$ es estable.
(*ii*) (*i*), (*ii*) y (*iii*) de 2.7, tomadas conjuntamente, son equivalentes a

$\sigma[\phi]$ es el más débil τ tan fuerte como σ tal que $\tau \Vdash \phi$.

2.9. COROLARIO

Sean $\psi_1, ..., \psi_n$, ϕ oraciones de $L_0^{\mathcal{A}}$. Entonces

$$\psi_1, ..., \psi_n \models_1 \phi \text{ syss } \psi_1, ..., \psi_n \models_2 \phi \text{ syss } \psi_1, ..., \psi_n \models_3 \phi$$

La siguiente observación se debe a Johan van Benthem (Cf. van Benthem [1989]). Lo que importa no es tanto la proposición como su demostración.

2.10. PROPOSICION

Sea ϕ una oración de $L_0^{\mathcal{A}}$. Entonces hay un subconjunto $[\![\phi]\!]$ de W tal que, para cualquier estado de información σ, $\sigma[\phi] = \sigma \cap [\![\phi]\!]$.

DEMOSTRACION

Hagamos $[\![\phi]\!] = \{w \in W \mid \{w\}[\phi] = \{w\}\}$ y supongamos que $\sigma = \{w_1,..., w_n\}$. Entonces

$$\sigma[\phi] = (\{w_1\} \cup ... \cup \{w_n\})[\phi].$$

Dados (*iii*) y (*iv*) de 2.7,

$$(\{w_1 \cup ... \cup \{w_n\})[\phi] = \{w_1\}[\phi] \cup ... \cup \{w_n\}[\phi].$$

De (*ii*) de 2.7 se sigue que para cada *i*, o bien $\{w_i\}[\phi] = \{w_i\}$ o bien $\{w_i\}[\phi] = \varnothing$ (el conjunto vacío). De aquí,

$$\{w_i\}[\phi] \cup ... \cup \{w_n\}[\phi] = \sigma \cap [\![\phi]\!].$$

La prueba muestra que en muchos casos el enfoque dinámico no tiene que ofrecer nada que no ofrezca ya el enfoque estático. Si resulta que la función de actualización [ ] es una operación distributiva y que aumenta monotónicamente (que es lo que dicen (*ii*), (*iii*) y (*iv*) de 2.7) sobre un retículo de estados de información (dados aquí mediante $< P(W), \subseteq >$), entonces puede uno asociar a cada oración ϕ un elemento fijo $[\![\phi]\!]$ del retículo —llamémosle la *proposición* expresada por ϕ— de tal modo que la actualización de un estado σ mediante ϕ equivale a efectuar la intersección. En tal caso, uno podría igualmente haber adoptado $[\![\,]\!]$ como noción básica.

La proposición 2.10 no vale para $L_1^{\mathcal{A}}$. Las oraciones de la forma *podría* ϕ no expresan una proposición. Cuando nos referimos a una oración ϕ de L_0, podemos hablar de "los mundos-ϕ", los mundos en los que es verdadera la proposición $[\![\phi]\!]$ expresada por ϕ. Las oraciones de L_0 son puramente descriptivas. Pero las oraciones de la forma *podría* ϕ no son descriptivas; no tendría sentido hablar de "los mundos-*podría* ϕ".

2.11. PROPOSICION

Sean $\psi_1,..., \psi_n$, ϕ oraciones de $L_0^{\mathcal{A}}$. Entonces $\psi_1,..., \psi_n \models_i \phi$ syss $\psi_1,..., \psi_n / \therefore \phi$ es válida en lógica clásica. (i = 1, 2, 3).

Generalizando este resultado podemos decir que en la semántica de actualización las oraciones descriptivas se rigen por la misma lógica que en la semántica de condiciones de verdad. Pero al cambiar la perspectiva se hace posible obtener una nueva manera de tratar con oraciones no-descriptivas.

OBSERVACION MARGINAL

En L_1 *podría* puede aparecer tan sólo como el operador más externo de una oración. Supongamos que se cambiara nuestra gramática de tal forma que *podría* pudiera darse en cualquier lugar y que todo lo demás siguiera igual.

(*i*) Es fácil apreciar que en tal caso la *Idempotencia* no seguiría estando válida.
(*ii*) ¿Existe quizás un argumento convincente en favor de la invalidez de la condición de *Idempotencia*?
(*iii*) ¿Existe quizás una forma elegante de restaurar la condición de *Idempotencia*? Después de todo, podría ocurrir que las condiciones de actualización dadas para "$\neg$", "$\wedge$" y "$\vee$" sean demasiado $L_0^{\mathcal{A}}$-específicas como para funcionar bien con oraciones que no expresen proposiciones.
(*iv*) ¿O deberíamos limitarnos a prohibir que *podría* aparezca en cualquier parte? Después de todo, *podría* expresa consistencia, una meta-noción.

Realmente, no sé qué posición adoptar aquí. Todo lo que sé es que las cosas se complican bastante si uno opta por (*iii*) y comienza a formular cláusulas como

$$\sigma\,[\phi \wedge \psi] \text{ es el menor } \tau \text{ tan fuerte como } \sigma \text{ tal que } \tau \models \phi \text{ y } \tau \models \psi;$$
$$\sigma\,[\neg p] \text{ es el menor } \tau \text{ tan fuerte como } \sigma \text{ tal que } \tau\,[\phi] = \mathbf{0}.$$

Por un lado, no habrá siempre un *único* τ que sea el más débil con las propiedades requeridas. (Considérese la oración $\neg$(*podría* $p \wedge$ *podría* $\neg p$).)

Referencias

Gärdenfors, P. 1984 "The Dynamics of Belief as a Basis of Logic", *British Journal for the Philosophy of Science*, 35, 1-10.

Groenendijk, J. y M. Stokhof 1989: "Dynamic Predicate Logic", *ITLI-Prepublication LP-89,02*, Universidad de Amsterdam.

Groeneveld, W. 1989: "A Dynamic Analysis of Paradoxical Sentences", tesis para el grado de master, Departamento de Filosofía, Universidad de Amsterdam.

Heim, I. 1982: *The Semantics of Definite and Indefinite Noun Phrases*, tesis para el grado de doctor, Universidad de Massachusetts, Amherst.

Kamp, J.A.W. 1981: "A Theory of Truth and Semantic Interpretation", en J. Groenendijk, T.M.V. Janssen y M. Stokhof, eds.: *Formal Methods in the Study of Language*, Mathematical Centre Tracts 135, Amsterdam, 277-322.

Makinson, D.: "General Theory of Cumulative Inference", en *Proceedings of the Second International Workshop on Non-monotic Reasoning*, Springer Lecture Notes on Computer Science.

STALNAKER, R. 1976: "Possible Worlds", *Noûs*, 10, 65-75.

STALNAKER, R. 1979: "Assertion", en P. Cole, ed.: *Syntax and Semantics 9 - Pragmatics*, Academic Press, New York, 315-332.

VAN BENTHEM, J.F.A.K. 1989: "Semantic Parallels in Natural Language and Computation", en H. Ebbinghaus et al., eds.: *Logic Colloquium '87*, Elsevier Science Publishers B.V. (North-Holland), Amsterdam, 331-375.

VELTMAN, F. 1991: "Defaults in Update Semantics", *ITLI-Prepublication* LP-91-01, Universidad de Amsterdam.

(Traducción de Juan José Acero)